Improving the Efficiency and Accuracy of Epileptic Seizure Detection using Machine Learning on EEG datasets

By Shubhan Anchan, Grade XI, Bombay International School

1. Origin of the Project

My interest in epilepsy detection was piqued by a personal experience in February 2024 when a relative afflicted with epilepsy spent 36 hours awaiting diagnosis after an epilepsy attach. The individual's suffering led me to explore the nuances of epilepsy detection, classification and treatment using electro-encephalography. I attended a 15 day workshop on "Nurturing AI Innovators" in November 2023 conducted by Prof. Pavlos Protopapas, Scientific Program Director, Institute for Applied Computational Science (IACS) at Harvard University during his visit to India, in collaboration with The Innovation Story, Mumbai. This learning led me to think hard about the potential application of machine learning to improve epilepsy detection and treatment practices.

I defined a project along with The Innovation Story to enable Neurologists to improve the efficacy of Epilepsy detection and treatment, by providing an Epilepsy Detection App that achieved the following:

- Reduces the time taken to detect epilepsy
- Improves the accuracy of epilepsy type detection (complex partial seizure, simple partial seizure, myoclonic seizure, absence seizure, etc)

by utilizing Machine Learning methodologies on Electroencephalography (EEG) data

2. Objective of the Project:

Epilepsy is a condition that affects the brain, causing repeated seizures. Epilepsy can start at any age and there are many different types. Some types last for a limited time, but for many people, epilepsy can be a life-long condition. 70 million people have epilepsy worldwide. There are more than 12 million persons with epilepsy (PWE) in India, which contributes to nearly one-sixth of the global burden. Epilepsy is the second most common and frequently encountered neurological condition that imposes heavy burden on individuals, families, and also on healthcare systems. People with epilepsy are at higher risk with osteoporosis, fractures as well as sleep deprivation and loss of memory. Research estimates cite 70% epilepsy patients could live seizure-free, if the seizure is diagnosed in time & classified accurately.

Epilepsy detection is primarily based on EEG Data. EEG Test Data is extremely voluminous for 1 test instance with data generation on 7 parameters from 21-23 electrical nodes placed on the brain over a 24-hour duration. This is a time-consuming affair requiring 15-20 minutes of analyses by experienced Neurologists. Classification of epilepsy is based on the nature and quantum of anomaly detected, which in turn determines treatment – with variability in evaluation across Neurologists based on expertise in EEG data assessment

The objective of this project is to classify different types of seizures (complex partial seizure, simple partial seizure, myoclonic seizure, absence seizure, etc) based on EEG recordings, more efficiently (>80% reduction in time) and accurately. The project seeks to enable remote diagnosis of epileptic seizures, improving the efficacy of epilepsy treatment in far-flung areas (Tier3 towns, rural areas) by providing access of EEG diagnosis outcomes for validation with expert neurologists (typically resident in Tier1 and 2 metros and towns).

The project seeks to provide the following

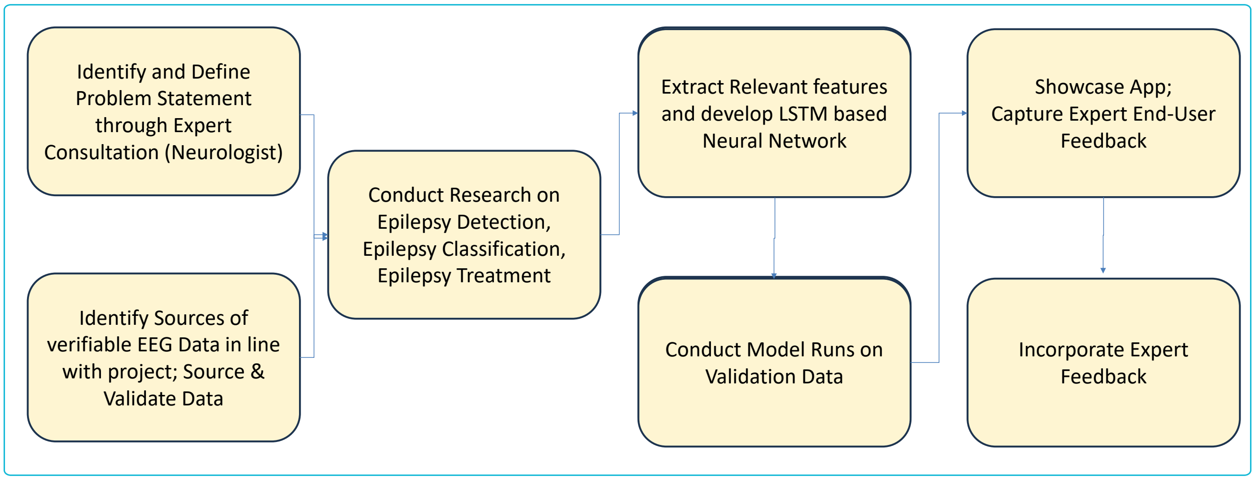
<u>Automate Seizure Classification</u>: Create a system that automatically classifies seizures from EEG signals, optimizing the time for manual checks by doctors

<u>Use Deep Learning</u>: Utilize deep learning techniques, specifically Long Short-Term Memory (LSTM) networks, to effectively identify patterns in the EEG data for better classification results.

<u>Build a User-Friendly Web Application</u>: Develop an easy-to-use web application where users can upload EEG data and receive quick results regarding seizure classification.

Shubhan Anchan,
Bombay International School

3. Research Methodology adopted for the Project



Defining the Problem to Solve

- **Neurologist Consultation**: I met with three neurologists at reputed Mumbai hospitals (Kokilaben Hospital, PD Hinduja Hospitals, NH-SRCC) to understand the challenges that the neurologist community face in diagnosing epilepsy
- The neurologists cited key challenges of time consumed in reading EEG datasets and accurately identifying & classifying epilepsy types – accentuated in remote areas of the country by the lack of access to experienced neurologists (causing extended delays in diagnosis)
- They cited the need to have a technology solution that reads EEG datasets, identifies anomalies and suggests an epilepsy classification (once identified).

Collection of EEG Data

- **Dataset**: The Temple University Hospital EEG Seizure Corpus (TUSZ) was identified and sourced (signing upto the necessary protocols). The same was used, containing EEG recordings and annotations showing when and what type of seizure occurred.

Preprocessing EEG Data

1. **Filtering the Signals**: EEG data can be noisy, so a filter was applied to keep only the frequencies relevant for brain activity, removing interference from muscle movements or eye blinks.
2. **Extracting Features**:

 Shubhan Anchan,
Bombay International School

- **Raw EEG Signals**: These were kept as-is to help the model learn the general patterns in the data.
- **Time-Based Features**: Basic statistical values were calculated, like the average signal value, variation in signal, and peak intensity.
- **Frequency-Based Features**: Specific frequency ranges (like Delta and Alpha waves) associated with different brain states were identified.
- **Wavelet Features**: A technique called the wavelet transform was used to capture details in both time and frequency, which is especially useful for detecting short bursts of seizure activity.

3. **Balancing Data**: Since some seizure types were rare, a technique called Synthetic Minority Over-sampling Technique (SMOTE) was used to create synthetic examples of the underrepresented types, so the model didn't focus only on common seizures.

Building Neural Network for EEG Detection and Classification

- **LSTM Model**: An LSTM (Long Short-Term Memory) neural network was used because it's well-suited for handling time-based data, helping to capture patterns in EEG signals over time. Additional adjustments, like dropout layers, were added to avoid overfitting (when a model learns too specifically from the training data and performs poorly on new data).

Running the LSTM model on EEG Datasets :

- **Web App**: A simple web app was created using Flask to allow users to upload EEG files, which the model then analyzes. The app provides an immediate classification of the seizure type and a visualization of when it occurred in the EEG data.
- Accuracy Testing: We tested the model on approximately anonymised EEG datasets for 29000 patients, covering the following
  - Complex Partial Seizures (cpsz)
  - Generalized Non-Specific Seizures (gnsz)
  - Focal Non-Specific Seizures (fnsz)
  - Tonic Clonic seizures (tcsz)
  - Absent seizures (absz)
  - Myoclonic seizures (mysz)
  - Tonic seizures (tnsz)
  - Simple Partial Seizures (spsz)
  - Atonic seizures (atsz)
  - Clonic seizures (cnsz)

  Testing yielded 96% accuracy with higher accuracy realized in those datasets with large data points.

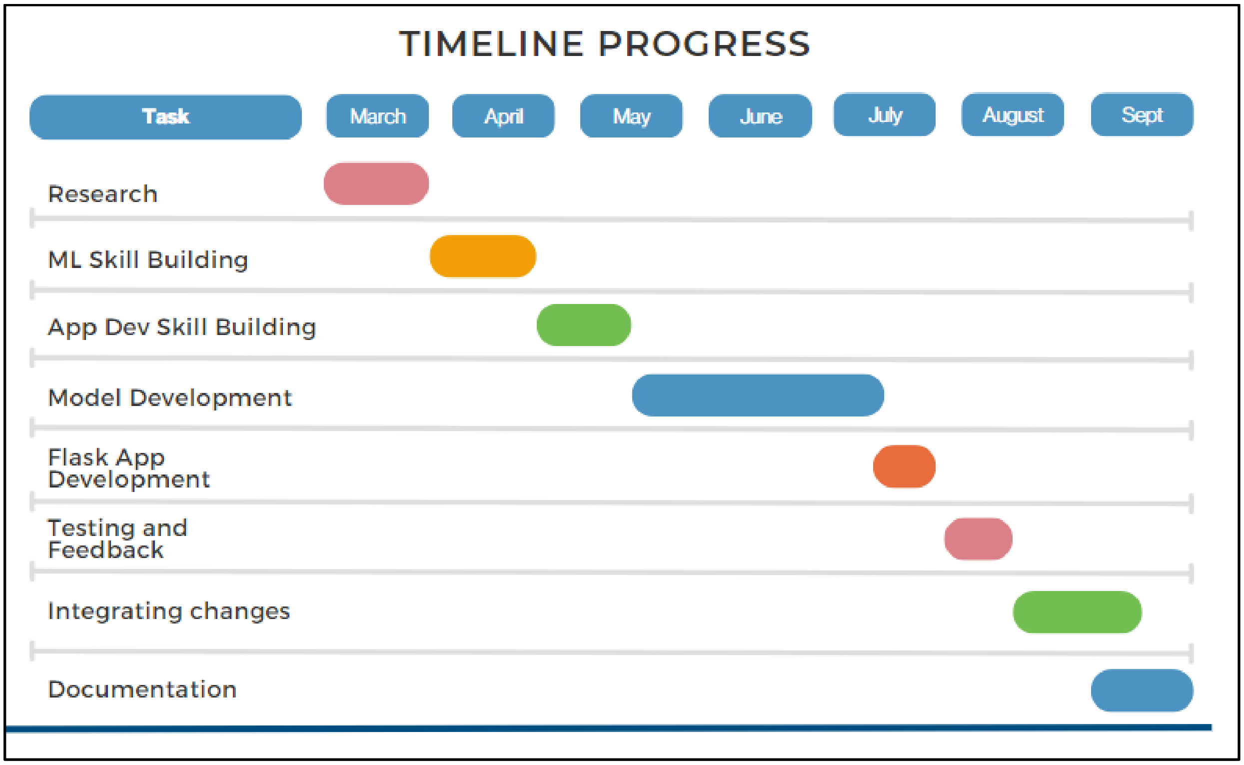
 Shubhan Anchan,
Bombay International School

Capture Expert Feedback

- **Neurologist Feedback**: I showcased the algorithm to the neurologists who appreciated the utility of the Epilepsy ML app. They offered suggestions on bringing the app to remote locations to be able to quickly capture first-pass diagnosis and classification which allowed them to review specific wavelets with anomalies and push forward on recommended treatment pathways.
- **Way Forward**: they offered suggestions for future development on capturing multiple EEG test configurations and visualizing multi-time interval EEG data sets superimposed to check for epilepsy progression and efficacy of treatment pathways.

4. Place of Project Research and Project Timeframe:

The project was executed with The Innovation Story, Dadar, Mumbai with neurologist consultations at Kokilaben Hospital, PD Hinduja Hospital and Narayana Hrudalaya-SRCC, Mumbai. I followed the timeline below:



Shubhan Anchan,
Bombay International School

5. Work Done

Research (March):

In March, I focused on finding studies and articles about detecting and classifying seizures using EEG signals. I looked into different methods used in previous research and checked what gaps my project could fill. I also searched for datasets that I could use to train my models later.

ML Skill Building (April):

In April, I worked on improving my machine learning skills. I learned about important concepts like feature extraction and different classification algorithms. I tried out simple models using Python libraries, which helped me understand how machine learning works and prepared me for my project.

App Development Skill Building (May):

In May, I shifted my focus to app development by learning how to create web applications with Flask. I followed tutorials to build basic apps and learned how to connect machine learning models to web frameworks. This training was important for making my machine learning model accessible to users through a web application.

Model Development (May - July):

From May to July, I worked on developing the seizure classification model. I processed the data, extracted features, and trained different machine learning models using EEG data. I chose to use a Long Short-Term Memory (LSTM) network because it can understand the sequence of data over time. I adjusted the model to improve its performance and checked its accuracy in classifying different types of seizures.

Flask App Development (July):

In July, I started building the Flask web application to integrate my seizure classification model. I set up the app's structure, created pages for uploading EEG data, and designed a user-friendly interface. I made sure the app was easy to use and added features to show the classification results clearly.

Testing and Feedback (August):

In August, I tested the web application and the model's performance thoroughly. I gathered feedback from the neurologists to find out what worked well and what needed improvement.

Integrating Changes (August - September):

From late August to September, I focused on making changes based on the feedback I received. I refined the model to improve its accuracy, optimized the app's code for better performance, and updated the app's layout to make it more user-friendly. I added new features, such as signal visualization.

Documentation (September):

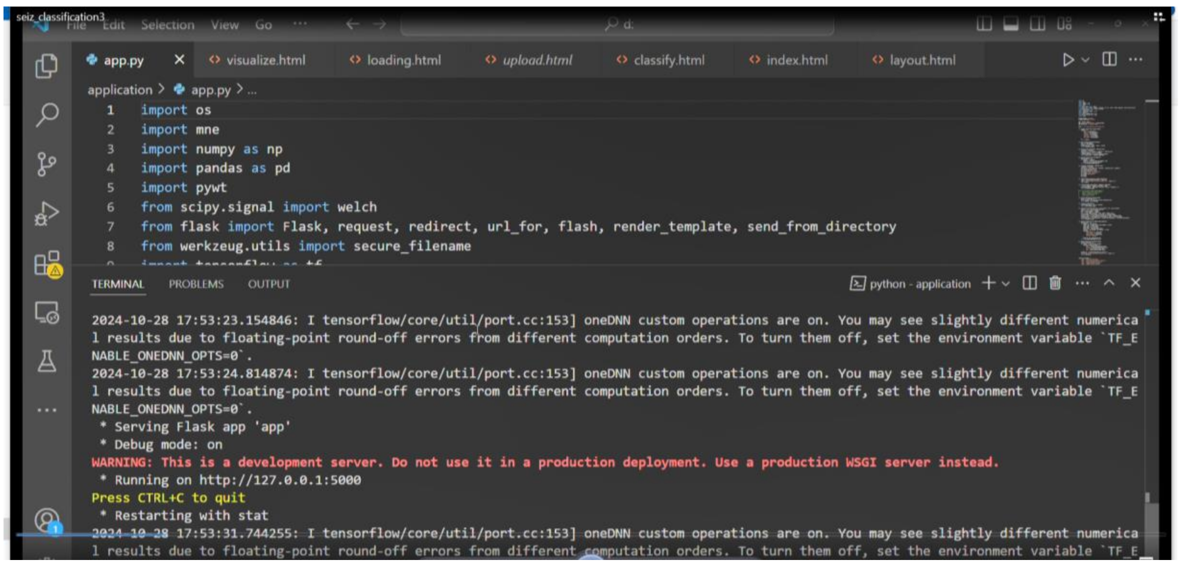
In September, I created detailed documentations for the project, explaining how the model works and how the Flask application was created.

6. Project Outcomes

The Epilepsy Detection app was able to identify and classify epilepsy instances (where applicable) in 45 seconds (as compared to current times of 12-15 minutes). Classification accuracy was recorded at 92% across epilepsy types, with 98% accuracy in the most statistically significant EEG datasets.

Exhibit below shows the Algorithm written for the Epilepsy Detection app, covering

- Visualization
- Data loading
- Identification and Classification
- Feedback



Shubhan Anchan,
Bombay International School

Exhibit below provides a snapshot of the User Interface for the ML app, covering Data Upload, Wavelet Identification for anomaly Classification and Classification results.

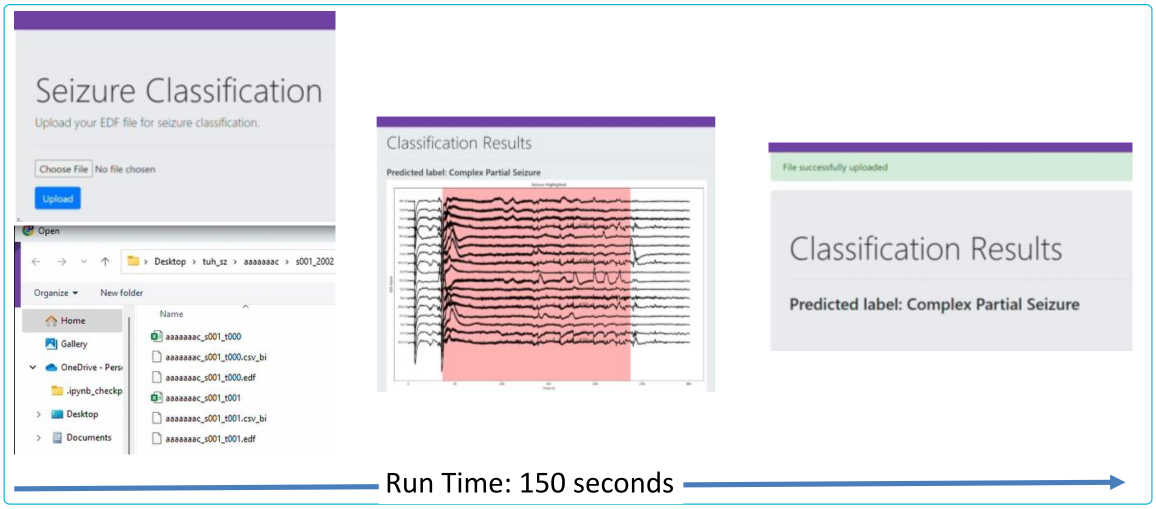


Exhibit below shows the testing outcomes on ~29,000 EEG datasets, yielding 92% accuracy of classification outcomes

Shubhan Anchan,
Bombay International School

7. Present Status of the Project & Future Plans:

Present Status:

- Successfully classified different seizure types using EEG Data from the Temple of University Seizure Corpus
- Utilized data preprocessing, feature extraction, and an LSTM Model for accurate classification
- Developed a user-friendly Flask application for real-time seizure type prediction.
- Reduced EEG analysis time significantly and improved consistency in seizure classification.

Future Plans:

Release 1 : Epilepsy Detection and Classification

- Incorporate multiple EEG setups towards normalized data input and consistent identification and classification based on node protocols
- Improve accuracy of the LSTM model by testing and providing classification feedback
- Showcase App to Indian Academy of Neurology and Capture their Feedback
- Showcase APP to National Health Mission for consideration under the Ayushman Bharat Digital Health Mission
- Refine MVP to showcase to Primary and Tertiary Health Providers, in remote areas

Release 2: Epilepsy Treatment Efficacy, tracking parameters over multiple test instances

- Develop algorithm for multiple test instance tracking (over time) of EEG parameters and epilepsy treatment pathways
- Test visualization with Expert Neurologists checking for accuracy of outcomes
- Execute on above scalability (defined for Release 1)

8. Financial Summary:

I have developed the app with the support of one AI-ML expert. During our research, we did not come across solutions that provided ML-based EEG dataset analyses. Current solutions are focused on providing neurologists to visualize EEG wavelet patterns DURING the manual diagnosis and classification. I propose to estimate the cost of the MVP as one of the next steps.

Shubhan Anchan,
Bombay International School

9. Acknowledgement of Mentors to the Project:

- Dr. Charu Sankhla, Faculty Neurologist & Movement Disorders Specialist, PD Hinduja Hospital, Mumbai
- Dr. Tushar Raut, Neurologist, Kokilaben Hospital, Mumbai
- Dr. Pradnya Gadgil, Neurologist, Narayan Hrudalaya SRCC, Mumbai
- Mr. Jaikumar Patel and Ms. Latifa Zaman, mentors for Python coding, The Innovation Story

10. Resources and References:

1. Temple University Hospital (TUH) Neural Engineering Data Consortium data and document repository
   Temple University EEG Corpus
2. Ali Shoeb. Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment. PhD Thesis, Massachusetts Institute of Technology, September 2009.
3. Guttag, J. (2010). Children's Hospital of Boston – Massachusetts Institute of Technology Scalp EEG Database (version 1.0.0). PhysioNet.
   CHB-MIT Scalp EEG Database v1.0.0
4. Enhanced Epileptic Seizure Detection through Wavelet-Based Analysis of EEG Signal Processing (Applied Sciences)
   by Sebastián Urbina Fredes, Ali Dehghan Firoozabadi, Pablo Adasme, David Zabala-Blanco, Pablo Palacios Játiva and Cesar Azurdia-Meza
   Applied Sciences 2024, 14(13), 5783; https://doi.org/10.3390/app14135783
5. Epileptic seizure detection combining power spectral density and high frequency oscillations (IJAMEC)
   by Rabia Tutuk and Reyhan Zengin
   https://dergipark.org.tr/en/download/article-file/2874211
6. Evaluating CNN Methods for Epileptic Seizure Type Classification Using EEG Data (IEEE Access)
   by Manuel J. Rivera, Javier Sanchis, Oscar Corcho,Miguel A. Teruel, Juan Trujillo,
   https://www.researchgate.net/publication/380952355_Evaluating_CNN_Methods_for_Epileptic_Seizure_Type_Classification_Using_EEG_Data
7. EEG epileptic seizure detection and classification based on dual-tree complex wavelet transform and machine learning algorithms by Itaf Ben Slimen, Larbi

Shubhan Anchan,
Bombay International School

Boubchir, Zouhair Mbarki, Hassene Seddik, The Journal of Biomedical Research, 2020 34(3): 151–161
https://pmc.ncbi.nlm.nih.gov/articles/PMC7324280/

8. EEG seizure detection: concepts, techniques, challenges, and future trends
By Ein Shoka, A.A., Dessouky, M.M., El-Sayed, A. et al. Multimedia Tools Applications 82, 42021–42051 (2023). https://doi.org/10.1007/s11042-023-15052-2
9. Epileptic Seizure Detection Technique Using EEG Signals with Mobile Application Development
By Lasefr, Zakareya, Khaled Elleithy, Ramasani Rakesh Reddy, Eman Abdelfattah, and Miad Faezipour. 2023. Applied Sciences 13, no. 17: 9571. https://doi.org/10.3390/app13179571
10. Epileptic seizure detection with deep EEG features by convolutional neural network and shallow classifiers
By Wei Zeng, Liangmin Shan, Bo Su and Shaoyi Du, Frontiers in Neuroscience, 22 May 2023, Section (Neuroprosthetics), Volume 17 - 2023 | https://doi.org/10.3389/fnins.2023.1145526
11. Automated Epileptic Seizure Detection Based on Wearable ECG and PPG in a Hospital Environment
By Vandecasteele, Kaat, Thomas De Cooman, Ying Gu, Evy Cleeren, Kasper Claes, Wim Van Paesschen, Sabine Van Huffel, and Borbála Hunyadi. 2017. Sensors 17, no. 10: 2338. https://doi.org/10.3390/s17102338
12. Wearable Epileptic Seizure Prediction System Based on Machine Learning Techniques Using ECG, PPG and EEG Signals
By David Zambrana-Vinaroz, Jose Maria Vicente-Samper, Juliana Manrique-Cordoba and Jose Maria Sabater-Navarro, Sensors 2022, 22, 9372. https://doi.org/10.3390/s22239372

 Shubhan Anchan,
Bombay International School